Computational Aspects of Fairness in Multi-Winner Voting

Simon Rey, Julian Chingoma and Jan Maly

June 2022 Project

Institute for Logic, Language and Computation
University of Amsterdam

Introduction

What did we do yesterday? We introduced several fairness criteria and discussed whether they could be achieved or not. This analysis left out several questions:

Can we efficiently verify that a fairness criteria is satisfied?

Can we efficiently find an outcome satisfying a given fairness criteria?

In today's lecture we will focus on those matters by investigating the computational aspects of fairness in multi-winner voting.

Computation Complexity

Complexity theory studies how hard it is to solve a problem.

- What do we mean by "solve"?

  ↪ We use an *abstract model of computation*—Turing machine—that gives a precise mathematical definition of what an algorithm is. Solving a problem then means defining a program for this abstract machine which answers the question.

- How do we measure the "hardness" of solving a problem?

  ↪ We count the *number of elementary steps* needed to run a program solving the problem on a Turing machine.

Decision Problems and P

We focus on *decision problems*, i.e., problems for which the answer is either Yes or No, and we group them by hardness.

DEFINITION: COMPLEXITY CLASS P

P is the class of all the decision problems for which there exists a Turing machine $\mathbb{M}$ answering the any instance x of the problem in time $\mathcal{O}(|x|^c)$, for a given $c \in \mathbb{N}$.

The Complexity Classes NP and coNP

Definition: Complexity Class NP

NP is the class of all the decision problems for which there exists a Turing machine $\mathbb{M}$ that can verify whether a potential solution (of polynomial size) of a problem instance x indeed is a solution for x, in time $\mathcal{O}(|x|^c)$, for a given $c \in \mathbb{N}$.

Definition: Complexity Class coNP

coNP is the class of all the decision problems for which there exists a Turing machine $\mathbb{M}$ that can verify whether a potential solution (of polynomial size) of a problem instance x *is not* a solution for x, in time $\mathcal{O}(|x|^c)$, for a given $c \in \mathbb{N}$.

Hardness

Membership to a complexity class can be seen as an upper-bound on the complexity of a problem: problems in P can be solved in polynomial time, problems in NP in non-deterministic polynomial time. What about lower-bound then?

A decision problem is *hard* for a complexity class if it is as hard to solve as any other problem in the class. To show that a problem is hard, we take another one from the class and prove that we can solve the latter by solving the former. This is called a *reduction*.

Problems that belong to a complexity class and that are hard for that class are called *complete for the class*.

Plan for the Lecture

We are now fully equipped to delve into the lecture, we will:

- Look at the complexity of checking if a committee satisfies concepts based on justified representation;
- Investigate the complexity of finding an outcome satisfying extended justified representation by:
  - Exploring the complexity of determining the winner under Thiele Rules (especially PAV, which satisfies EJR);
  - Presenting a new rule that satisfies EJR... in polynomial time.

Shall we?

1. Justified Representation

Let's Start Easily

Definition: Cohesive Group

For $\ell \geq 1$, a group of agents $N' \subseteq N$ is ℓ-cohesive if $|N'| \geq \ell \times {}^n/_k$, and $|\bigcap_{i \in N'} A_i| \geq \ell$.

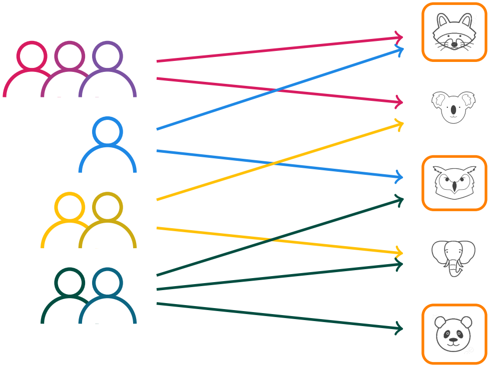
Definition: Justified Representation

A committee C of size k satisfies Justified Representation (JR) if for every 1-cohesive group $N' \subseteq N$, there exists $i \in N'$ such that $A_i \cap C \neq \emptyset$.

↳ What is the complexity of checking whether a committee C satisfies JR?

Checking Justified Representation

$k = 3$



Number of Approvers	Cohesiveness	Number of Unsatisfied Approvers
$4 \geq 8/3$	✓	$0 < 8/3$
$5 \geq 8/3$	✓	$2 < 8/3$
$3 \geq 8/3$	✓	$0 < 8/3$
$4 \geq 8/3$	✓	$2 < 8/3$
$2 < 8/3$	✗	—

Algorithm: For every $c \in C$, check if $N' = \{i \in N \mid c \in A_i\}$ is 1-cohesive, and if so, check if less than n/k agents from N' have no representative.

Now Comes the Hard Part

Definition: Cohesive Group

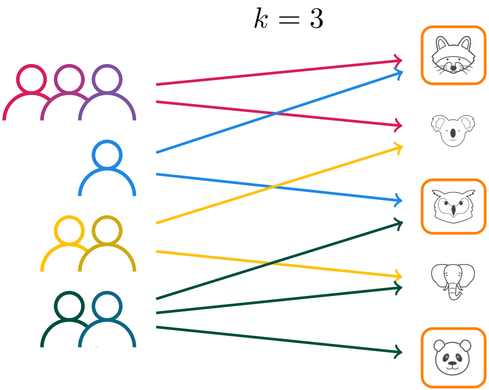
For $\ell \geq 1$, a group of agents $N' \subseteq N$ is ℓ-cohesive if $|N'| \geq \ell \times n/k$, and $|\bigcap_{i \in N'} A_i| \geq \ell$.

Definition: Extended Justified Representation

A committee C of size k satisfies Extended Justified Representation (EJR) if for every $\ell \in \{1, \ldots, m\}$ and every ℓ-cohesive group $N' \subseteq N$, there exists $i \in N'$ such that $|A_i \cap C| \geq \ell$.

↳ What is the complexity of checking whether a committee C satisfies EJR?

Checking Extended Justified Representation



PROPOSITION:

Checking whether a committee C satisfies EJR is a coNP-complete problem (even checking for the existence of an ℓ-cohesive group is an NP-hard problem).

Hardness of a Fairness Criteria or Hardness of a Rule?

Checking whether a committee satisfies EJR is coNP-complete. Does that imply rules finding EJR committee are all hard to compute?

Spoiler alert: No!

↳ Let's look into computing the outcome of PAV since we know it satisfies EJR.

2. Thiele Rules and Proportional Approval Voting

Thiele Rules

Definition: Thiele Rules

Given a profile $\boldsymbol{A} = (A_1, \ldots, A_n)$, the w-Thiele rule associated with the weight function $w : \mathbb{N}_{>0} \rightarrow \mathbb{R}_{\geq 0}$ such that:

$$\arg\max_{W \in \mathcal{P}_k(C)} \sum_{i \in N} \sum_{j=1}^{|W \cap A_i|} w(j).$$

Not that the definition from the book differs as they works with weight functions $w'(|W \cap A_i|) = \sum_{j=1}^{|W \cap A_i|} w(j)$.

- AV is the w_{AV}-Thiele rule where $w_{AV}(j) = 1$ for all j. *Easy!*
- CC is the w_{CC}-Thiele rule where $w_{CC} = \begin{cases} 1 & \text{if } j = 1 \\ 0 & \text{otherwise} \end{cases}$. *Hard! (Remember the CC exam)*
- PAV is the w_{PAV}-Thiele rule where $w_{PAV} = 1/j$ for all j. *Hard!*

↳ For which of these rules (if any) would it be hard to compute the winning committee?

What Does Hardness of a Rule Mean?

We need to consider decision problems, but determining the outcome of a rule is not a decision problem. We will use the following decision problem:

w-Thiele Winner Determination	
Instance:	A set of candidates C, a profile $\boldsymbol{A} = (A_1, \ldots, A_n)$ over C, an integer $k \in \mathbb{N}$ and a threshold $s \in \mathbb{R}$.
Question:	Is there a committee $W \in \mathcal{P}_k(C)$ such that: $\sum_{i \in N} \sum_{j=1}^{\lvert W \cap A_i \rvert} w(j) \geq s$?

↳ Is this a good decision problem?

Proposition:

If w-Thiele Winner Determination is NP-hard then there exists no polynomial algorithm computing the outcome of the w-Thiele rule, unless P = NP.

Hardness of Winner Determination for Thiele Rules

Proposition:

For every weight function $w : \mathbb{N}_{>0} \rightarrow \mathbb{R}$ for which there exists p such that $w(p) > w(p+1)$, w-Thiele Winner Determination is NP-complete.

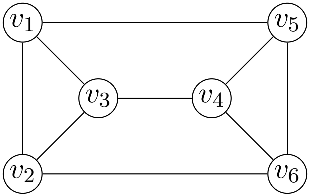
↳ This result would apply for w_{CC}, w_{PAV} but not w_{AV} (and also not for $w(j) = 2j - 1$ for instance, i.e., the square of the number of representatives).

Actually, w_{AV} is the only non-increasing weight function for which this result does not apply.

NP-hardness of Thiele Rules – Cubic Graph Vertex Cover

Cubic Graph Vertex Cover Problem: Given a graph $G = (V, E)$, every vertex has degree exactly 3 and an integer K, does G has a vertex cover of size K?

- The degree of a vertex is the number of edges involving that vertex.
- A vertex cover is a subset of vertices that contains at least one end-point of every edge.

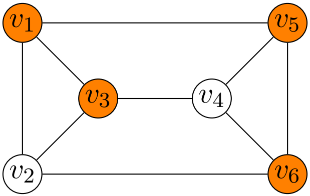


NP-hardness of Thiele Rules – Cubic Graph Vertex Cover

Cubic Graph Vertex Cover Problem: Given a graph $G = (V, E)$, every vertex has degree exactly 3 and an integer K, does G has a vertex cover of size K?

- The degree of a vertex is the number of edges involving that vertex.
- A vertex cover is a subset of vertices that contains at least one end-point of every edge.

$K = 4$

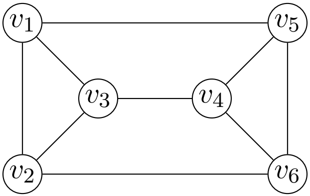


NP-hardness of Thiele Rules – Cubic Graph Vertex Cover

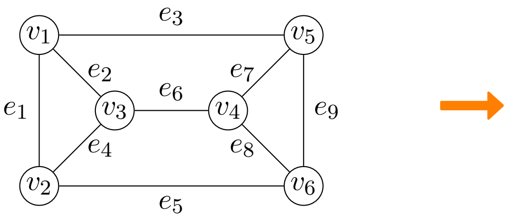
Cubic Graph Vertex Cover Problem: Given a graph $G = (V, E)$, every vertex has degree exactly 3 and an integer K, does G has a vertex cover of size K?

- The degree of a vertex is the number of edges involving that vertex.
- A vertex cover is a subset of vertices that contains at least one end-point of every edge.

$K = 3$



Reminder: p is the position such that $w(p) > w(p+1)$.



$$N = \{e_1, \dots, e_9\}$$
$$C = \{v_1, \dots, v_6\} \cup \{b_1, \dots, b_{p-1}\}$$
$$A_1 = \{b_1, \dots, b_{p-1}\} \cup \{v_1, v_2\}$$
$$A_2 = \{b_1, \dots, b_{p-1}\} \cup \{v_1, v_3\}$$
$$\vdots$$
$$A_9 = \{b_1, \dots, b_{p-1}\} \cup \{v_5, v_6\}$$
$$k = K + p - 1$$
$$s = n \sum_{j=1}^{p} w(j) + (3K - n) w(p+1)$$

Note: This can be done in time polynomial in the size of G.

Claim: There is a vertex cover of size K if and only if there exists a committee whose w-Thiele score is at most s.

$\mathtt{NP}$-hardness of Thiele Rules – Left-to-Right

$$N = E \qquad C = V \cup \{b_1, \dots, b_{p-1}\} \qquad A_i = \{b_1, \dots, b_{p-1}\} \cup \{e_i^1, e_i^2\} \qquad k = K + p - 1$$

$$s = n \sum_{j=1}^{p} w(j) + (3K - n)w(p+1)$$

Consider a vertex cover V' of size K. Let $W = V' \cup \{b_1, \dots, b_{p-1}\}$ be a committee.

For any agent e_i, $|A_i \cap W| \geq p$, indeed: the $p - 1$ dummy candidate and one of e_i^1 or e_i^2.

↳ Those contribute $n \sum_{j=1}^{p} w(j)$ to the score of W.

A vertex cover of size K covers exactly n edges, thus $3K - n$ edges are covered twice, i.e., exactly $3K - n$ agents have $p + 1$ representatives in W.

↳ Those contribute $(3K - n)w(p + 1)$ to the score of W.

The total score of W is thus $n \sum_{j=1}^{p} w(j) + (3K - n)w(p + 1) = s$. ■

NP-hardness of Thiele Rules – Right-to-Left

$$N = E \qquad C = V \cup \{b_1, \ldots, b_{p-1}\} \qquad A_i = \{b_1, \ldots, b_{p-1}\} \cup \{e_i^1, e_i^2\} \qquad k = K + p - 1$$

$$s = n \sum_{j=1}^{p} w(j) + (3K - n)w(p+1)$$

Consider set W of $K + p - 1$ candidates and assume that it does not correspond to a vertex cover, i.e., there is $e_i \in N$ such that $\{e_i^1, e_i^2\} \cap W = \emptyset$. We show that W's score is less than s.

To maximize the score, you must pick $\{b_1, \ldots, b_{p-1}\}$, those contribute $n \sum_{j=1}^{p-1} w(j)$ to W's score.

K extra candidates have to be selected, these cover $3K$ agents. At most $n - 1$ agents are covered once (from our hypothesis on W), and thus at most $3K - (n-1)$ twice.

↳ Those contribute $(n-1)w(p) + (3K - (n-1))w(p+1)$ to the score of W.

$$n \sum_{j=1}^{p-1} w(j) + (n-1)w(p) + (3K - n + 1)(p+1) = s - w(p) + w(p+1) < s. \quad \blacksquare$$

Computing PAV winners

The previous result shows that we cannot hope to compute the outcome of PAV in polynomial time (unless $\mathtt{P} = \mathtt{NP}$). Several options are possible if you still want to compute PAV winners:

- Using integer linear programming solvers (as presented in the book);
- Considering fixed-parameter tractable algorithms (for instance, only exponential in m);
- Looking into approximation algorithms (there is a 0.7965-approximation algorithm for PAV).

All of that is nice, but does not answer what seems to have become our main question:

> Can we compute committee satisfying EJR in polynomial time?

3. The Method of Equal Share

The Method of Equal Share (MES) (Formerly Known as Rule X)

This is a sequential rule in which agents "buy" candidates. Every agent is initially given a *budget of 1*. To buy a candidate, agents have to pay a *price of* n/k.

MES goes in rounds, filling in a committee W, initially empty. Let $b_i(t)$ be the budget of agent $i \in N$ at round t. For any $\alpha \in \mathbb{R}$, a candidate $c \in C \setminus W$ is *α-affordable* at round t if:

$$\sum_{\substack{i \in N \\ c \in A_i}} \min(\alpha, b_i(t)) \geq n/k.$$

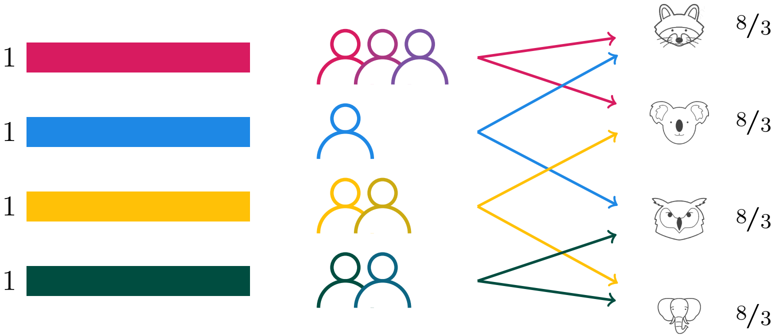
If at least one candidate is α-affordable for some $\alpha \in \mathbb{R}$, then MES selects a candidate $c \in C$ affordable for the smallest α (arbitrary tie-breaking). c is then added to W. The budget of every agent i approving of c is reduced by $\min(\alpha, b_i(t))$. A new round then starts.

Whenever no candidate is α-affordable, if $|W| < k$, we add arbitrary candidates to W to ensure that $|W| = k$. MES then terminates and returns W.

Remark: Do you see why we may need to add extra candidates?

An Example for MES

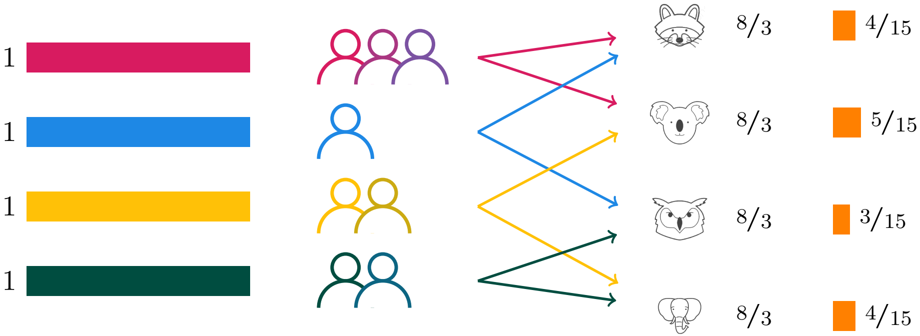
$$k = 3$$

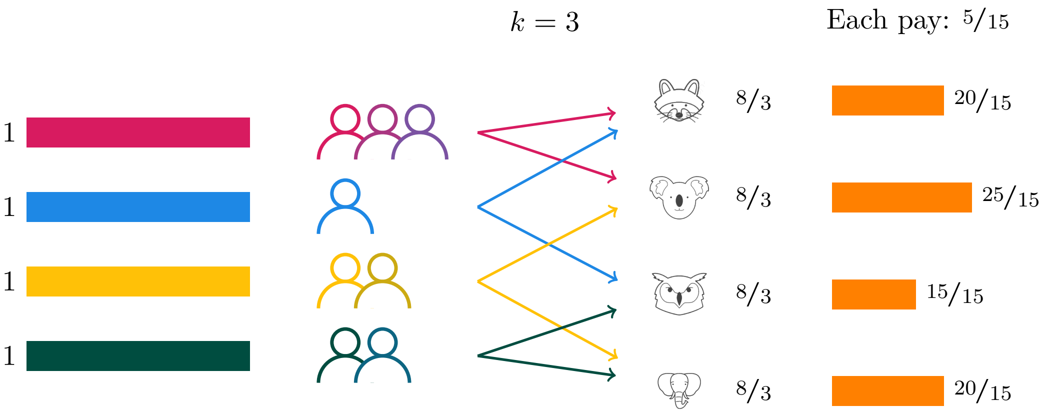


An Example for MES

$k = 3$

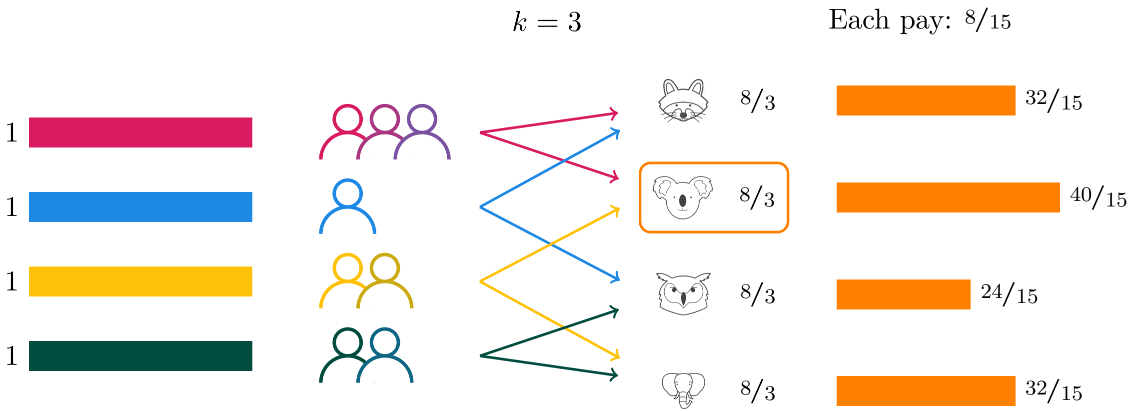
Each pay: $1/15$





$k = 3$
Each pay: 5/15
1
1
1
1
8/3
8/3
8/3
8/3
20/15
25/15
15/15
20/15

An Example for MES



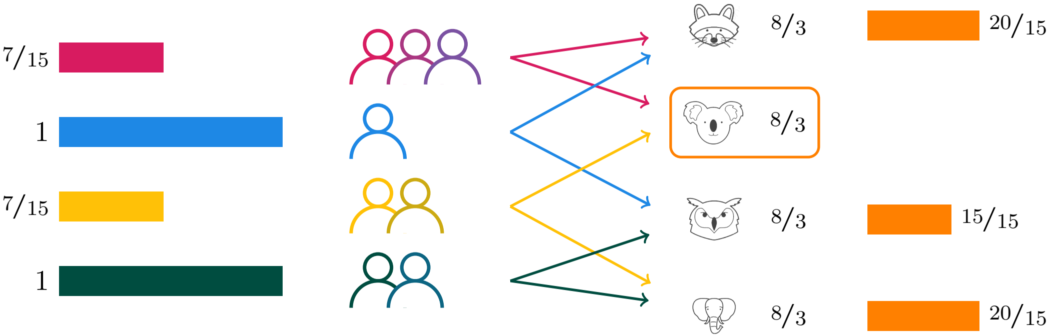
An Example for MES

$k = 3$

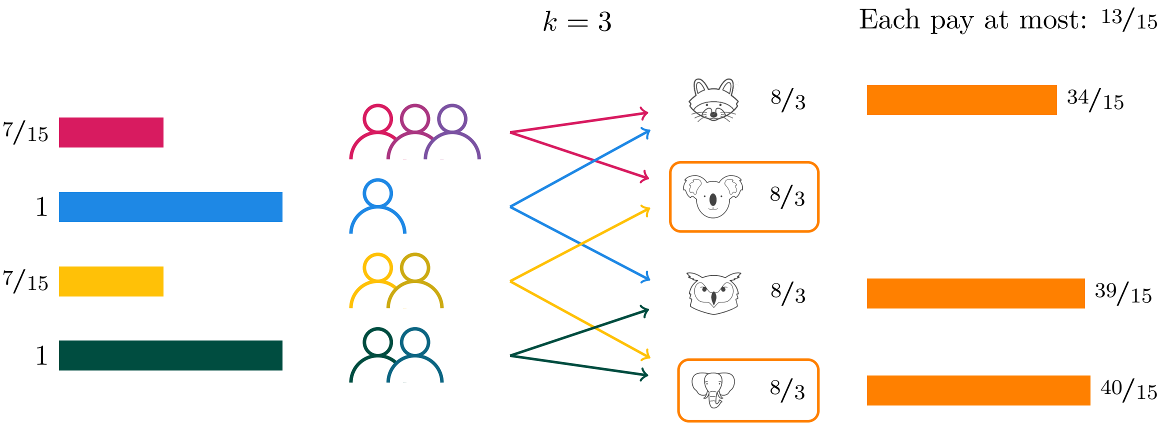
An Example for MES

$k = 3$

Each pay: $5/15$



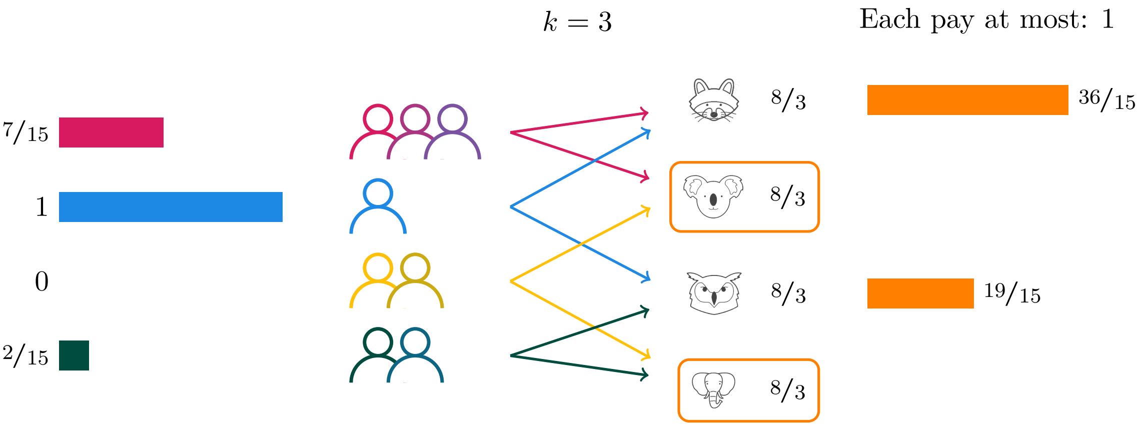
An Example for MES



An Example for MES

$k = 3$

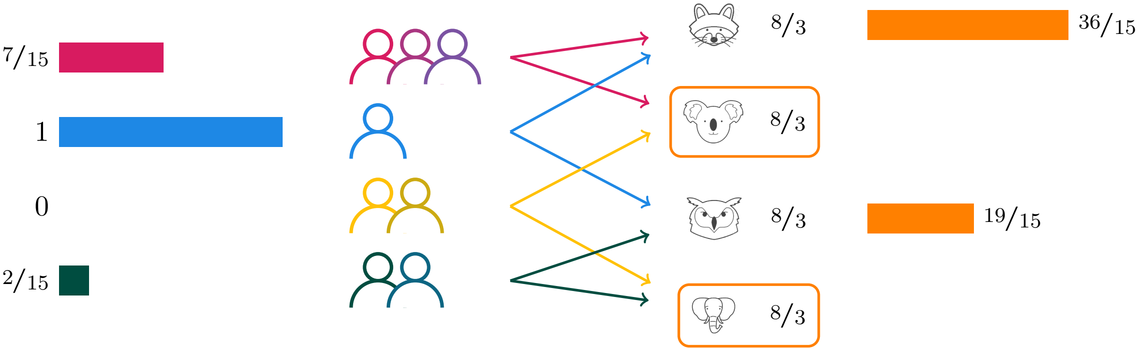
An Example for MES



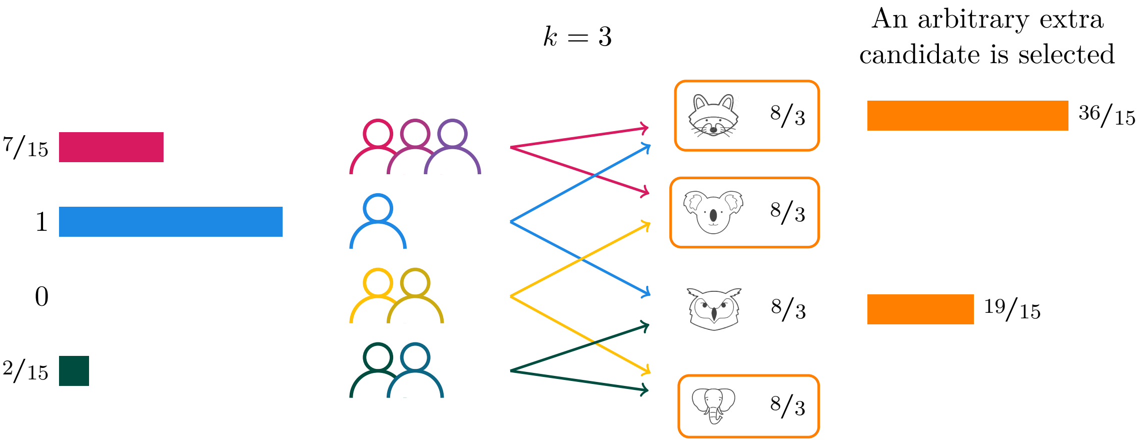
An Example for MES

$k = 3$

No candidate is affordable



An Example for MES



Efficiently Computing the Outcome of MES

Can we compute an outcome of MES efficiently? Yes!

We can compute the smallest α_c (if it exists) such that a given candidate $c \in C \setminus W$ is α_c-affordable by solving the following linear program:

$$
\begin{aligned}
\textbf{minimize} \quad & \alpha_c \\
\textbf{subject to:} \quad & \sum_{\substack{i \in N \\ c \in A_i}} \gamma_i \geq n/k \\
& 0 \leq \gamma_i \leq \alpha_c \qquad \forall i \in N, c \in A_i \\
& 0 \leq \gamma_i \leq b_i(t) \qquad \forall i \in N, c \in A_i \\
\textbf{where} \quad & \gamma_i \text{ is the contribution of agent } i \text{ to buy } c
\end{aligned}
$$

Good news: This is a continuous linear program (all variables are rational), we can solve it in polynomial time. We can thus compute an outcome of MES efficiently!

MES and Extended Justified Representation

Does MES provide extended justified representation? Yes!

Let W be returned by MES, $p = {}^n/_k$, and let N' be ℓ-cohesive such that $\forall i \in N'$, $|A_i \cap W| \leq \ell - 1$.

There must be $i \in N'$ such that $b_i(t^{end}) < {}^p/_{|N'|}$ (agents in N' could have bought something otherwise). Moreover, by cohesiveness ${}^p/_{|N'|} \leq {}^1/_\ell$. Since i has at most $\ell - 1$ representative in W, there exist $c \in W$ for which i paid strictly more than $\dfrac{1 - {}^p/_{|N'|}}{\ell - 1} \geq \dfrac{1 - {}^1/_\ell}{\ell - 1} = {}^1/_\ell$.

Let $c^\star$ be the first candidate selected by MES (at round $t^\star$) such that some voter from N' paid more than ${}^1/_\ell$ for. Thanks to the above, such a $c^\star$ exists. $c^\star$ is thus not ${}^1/_\ell$ affordable.

At $t^\star$, every voters in N' have at most $\ell - 1$ representatives and paid at most ${}^1/_\ell$ for them.

➥ Their leftover budget is at least $1 - (\ell - 1){}^1/_\ell = {}^1/_\ell$.

By cohesiveness, we have ${}^{|N'|}/_\ell \geq p$. There exists thus c' approved by every agent in N' which is ${}^1/_\ell$-affordable. MES should thus have selected c' and not $c^\star$. The contradiction is set. ∎

4. Conclusion

Wrap Up

Today we have investigating computational issues...

...Starting with the complexity of verifying whether a committee provide justified representation (easy) and extended justified representation (hard)...

...Moved to the problem of computing the outcome of Thiele rules, in particular PAV (hard)...

...Concluded on the existence of MES, a rules satisfying EJR and easy to compute.

Tomorrow, Jan will present you an extension of the framework in which candidates can have different cost: *participatory budgeting*.

Choosing Your Presentation

Tomorrow you will have to choose a topic for the presentations next week. Here are the topics:

1. More on Apportionment
2. Proportional Justified Representation and Phragmén's Rule
3. Priceability
4. Price of Fairness
5. Impossibility of Achieving Proportionality and Strategyproofness Simultaneously
6. Fairness in Perpetual Voting
7. The Method of Equal Share for Participatory Budgeting
8. Defining Fairness in Terms of Effort